Scottish Government
Riaghaltas na h-Alba
gov.scot

Evaluability Assessment of Scotland's Baby Box - Report to the Scottish Government

CHILDREN, EDUCATION AND SKILLS

social research

Evaluability Assessment of Scotland's Baby Box

Executive summary

Background

This report presents the Evaluability Assessment (EA) of Scotland's Baby Box scheme. The Baby Box scheme was piloted in Clackmannanshire and Orkney, Scotland from January to March 2017. It was then fully rolled out across Scotland in August 2017. Baby Boxes are new to Scotland and there is little available evaluation evidence to date on their impact from other countries. The Scottish Government is committed to evaluating the impact of the Baby Box initiative in Scotland and requested an EA to seek the advice of independent experts in identifying priorities for, and approaches to, evaluation.

The Evaluability Assessment process

Evaluability Assessment is a systematic, collaborative approach to the planning of an evaluation that involves engaging stakeholders, clarifying intervention goals, developing a theory of change or a logic model and deciding whether and how a useful evaluation could be carried out at a reasonable cost. The EA was conducted over the course of three workshops, between January and March 2018. Stakeholders who participated in the workshops were from the Midwifery profession, Scottish Government, third sector and charitable organisations, and academic researchers. The EA was conducted by the Evaluability Assessment Collaborative (EAC) on behalf of the Scottish Government.

Evaluation options

Stakeholders agreed that the evaluation should include both outcome and process evaluation elements, to identify the effect of the Baby Box and to explore the processes and mechanisms by which effects were achieved. It was agreed that the evaluation should focus on short to medium term outcomes, which could be more readily attributed to the Baby Box. It was recognised that the Baby Box may have effects on important rare or longer term child health, wellbeing and development outcomes, but these were likely also to be impacted by a wide range of other policies and external factors, such as the Universal Health Visiting Pathway and Best Start Grant, making it problematic to identify and separate the impact of the Baby Box, as is the case with evaluation of other specific initiatives. It was agreed that rare or longer term outcomes should be monitored to identify possible effects on health and wellbeing or unintended consequences that might indicate a need for further research.

It was agreed that a process evaluation should be carried out. This might use qualitative interviews with practitioners and Baby Box recipients to examine

implementation, mechanisms of impact, and contextual factors. This could be supplemented by case note reviews and a quantitative survey of recipients.

Evaluation of outcomes could potentially use data collected through the Child Health Systems Programme (CHSP) Pre-School system of assessments carried out by Health Visitors. Assessments are carried out at specified intervals following the birth of a child, and collect data on a range of relevant outcomes, including infant feeding, sleeping position, development and referrals to other services. The system is intended to ensure consistent recording of outcomes, but its primary purpose is to support the delivery of services. Therefore, validation work would be required to assess its suitability for evaluation uses.

Alternatively, primary data could be collected from a prospective cohort of births in Scotland and a comparison area (or areas) in England or Wales. This would allow much greater flexibility in deciding what data could be collected but would involve substantial methodological challenges and could be considerably more expensive than the other options.

In the absence of previous evaluations of Baby Box interventions delivered on a whole population basis, and the difficulties identified with the options for evaluating impacts, a phased approach to commissioning evaluation may be advisable. A substantial initial phase of feasibility testing and development work would be required prior to final decisions about the design of a substantive process and outcome evaluation.

Chapter 1: Background and Policy Context

Scotland's Baby Box scheme is part of a range of Scottish Government policies and initiatives intended to achieve the best possible outcomes for children and families in Scotland. Unlike initiatives such as the Healthy Start vouchers provided for low-income families, Baby Boxes are universally available for all babies born in Scotland, regardless of socioeconomic status or income. They provide families with essential items such as digital thermometers, clothes, books, a changing mat, blankets, and other items needed in the baby's first few months of life.

The Boxes are intended to increase parents' understanding and practice of positive behaviours such as safe sleeping, attachment, and breastfeeding. They also provide health professionals with an opportunity to engage families with a wide range of health services, and to initiate conversations around positive and risk behaviours. By providing parents with the necessary items and information, the Baby Box could positively impact a range of health and wellbeing outcomes for the child.

The evidence base

Evidence for the effectiveness of Baby Boxes and similar interventions for improving child health and development in countries comparable to Scotland is scarce. Finland introduced Baby Boxes in 1938 to address concerns with high infant mortality rates and decreasing birth rates (Embassy of Finland, 2017; Gissler and Hakulinen, 2017). The Baby Box, or maternity package, was initially available only for low-income parents, but since 1949 has been available to all mothers in Finland. Perinatal and infant outcomes in Finland have improved since the introduction of the Boxes, and are now among the best in the world (OECD 2018), but the impact of the scheme has not been formally evaluated.

Researchers acknowledge that any impact from the Boxes cannot be separated from other factors such as increased use of antibiotics, other improvements in healthcare, and the development of the wider welfare system (Gissler and Hakulinen, 2017).

Currently, Finland and Scotland are the only countries that operate universal, non-commercial Baby Box schemes. There are a number of schemes in other countries that share some features of Scotland's Baby Box, but they vary in accessibility, contents and messaging, and there have been few formal evaluations of their impact on health, wellbeing, and developmental outcomes for children and families. For example in New Zealand, a trial comparing a traditional woven flax bassinet that could be used in the parental bed with a conventional bassinet reported no increase in behaviours or markers of risk associated with sudden infant death, and some advantages, such as an increase in sustained breastfeeding (Tipene-Leach et al., 2014; Baddock et al., 2017). The Welcome to Parenthood scheme in Alberta, Canada provides parents having a first child with a Baby Box as part of a wider

package of mentoring support, and is being evaluated by researchers at the University of Calgary[1]. Results are not yet available.

Policy implementation

The Scottish Government has commissioned APS Group (Scotland) to provide the Baby Boxes. The four year contract is valued at £35.3 million for the delivery of the Baby Box.

Registration for full roll-out of the Box began on 15 June 2017 in preparation for nationwide distribution starting on 15 August 2017. Registration for a Scottish Baby Box has been made straightforward to encourage high take-up rates. Midwives administer and help complete registration cards with all pregnant women during the 20 to 24 week antenatal appointment. Once completed, registration cards are sent to APS Group, which operates a helpline for parents to contact should they have any enquiries once they have registered for a Box. Baby Boxes are delivered to parents between 32 and 36 weeks of pregnancy, and can be cancelled at any time leading up to delivery of the Box.

Previous research in Scotland

The Children and Families Directorate of the Scottish Government has commissioned three pieces of research during the development and early implementation of the Baby Box. The first was a qualitative study and an online survey to elicit parents' opinions of the contents and concept of the Baby Box. The findings informed the design and contents of the Baby Box prior to piloting.

The second research project was carried out on the Baby Box pilot, which took place from January to March of 2017. The pilot involved the delivery of 160 boxes to families and was conducted in two locations – Clackmannanshire and Orkney – chosen for their varying geographic and demographic characteristics. Interviews with families and health professionals were conducted at the pilot sites to obtain information on the use of the Box and its contents, the process for registering for and receiving the Box, and the overall response to the Baby Box scheme. Based on findings from the pilot research, the Scottish Government made several changes to the scheme prior to full roll-out.

The third research project took place after full roll-out of the Baby Box. A telephone survey was conducted with a sample of parents who had received the Baby Box to obtain their views on the contents of the Box, the nature and level of information included (e.g. on safe sleeping and breastfeeding), communications about the scheme and use of the associated Parent Club website. Overall satisfaction rates were very high, with 100% of parents stating they were 'very' or 'quite' satisfied with the overall quality of the Box and contents. Information from this study was used to inform year 2 procurement of the contents of the Box.

[1] http://www.ucalgary.ca/brightfromthestart/projects/w2p-ab.

The case for evaluation

Research carried out before and during implementation suggests that the Baby Box scheme has been well-received by parents, but evidence on the impacts of Baby Box-type interventions remains scarce. The Scottish Government is committed to evaluating the impact of the Baby Box on children and families in Scotland. To inform decisions about the scope and design of evaluation, and involve expert stakeholders reaching these decisions, the Children and Families Directorate commissioned an Evaluability Assessment. Chapter 2 describes the Evaluability Assessment process, Chapter 3 sets out and appraises evaluation options, and Chapter 4 provides a concluding discussion.

Chapter 2: The Evaluability Assessment process

Evaluability Assessment (EA) is a systematic, collaborative approach to the planning of an evaluation project. It involves engaging stakeholders, clarifying intervention goals, developing a theory of change (logic model) and deciding whether and how a useful evaluation could be carried out at a reasonable cost.

This EA was conducted on behalf of the Scottish Government by the Evaluability Assessment Collaborative (EAC). The EAC is a consortium of researchers from the Medical Research Council/ Chief Scientist Office Social and Public Health Sciences Unit, University of Glasgow and the Scottish Collaboration for Public Health Research and Policy, University of Edinburgh. The EA comprised three workshops, held between January and March 2018.

A wide range of stakeholders were identified and invited to take part in the workshops. They included staff from the Scottish Government, Midwifery, third sector and charitable organisations, and academic experts in safe sleeping and infant mortality. The full list of the workshop participants is provided in Appendix 1.

Workshop 1 (24 January 2018)

The first workshop began with an introduction to the EA process by the EAC. This was followed by a brief policy overview of the Baby Box by Scottish Government policy leads, which set out the rationale and purpose of the Baby Box initiative. The EAC then presented the logic model, which had been developed as part of the pilot evaluation of the Baby Box. Stakeholders agreed that it would be more useful to review and revise the existing logic model, rather than develop a new one, to reflect current operations and learning.

Stakeholders were divided into three sub-groups and were asked to discuss whether all the important inputs, activities and outcomes were captured in the logic model and to suggest other relevant items. The findings from each sub-group were then considered by the main group. The main points that emerged are outlined below.

Short term outcomes

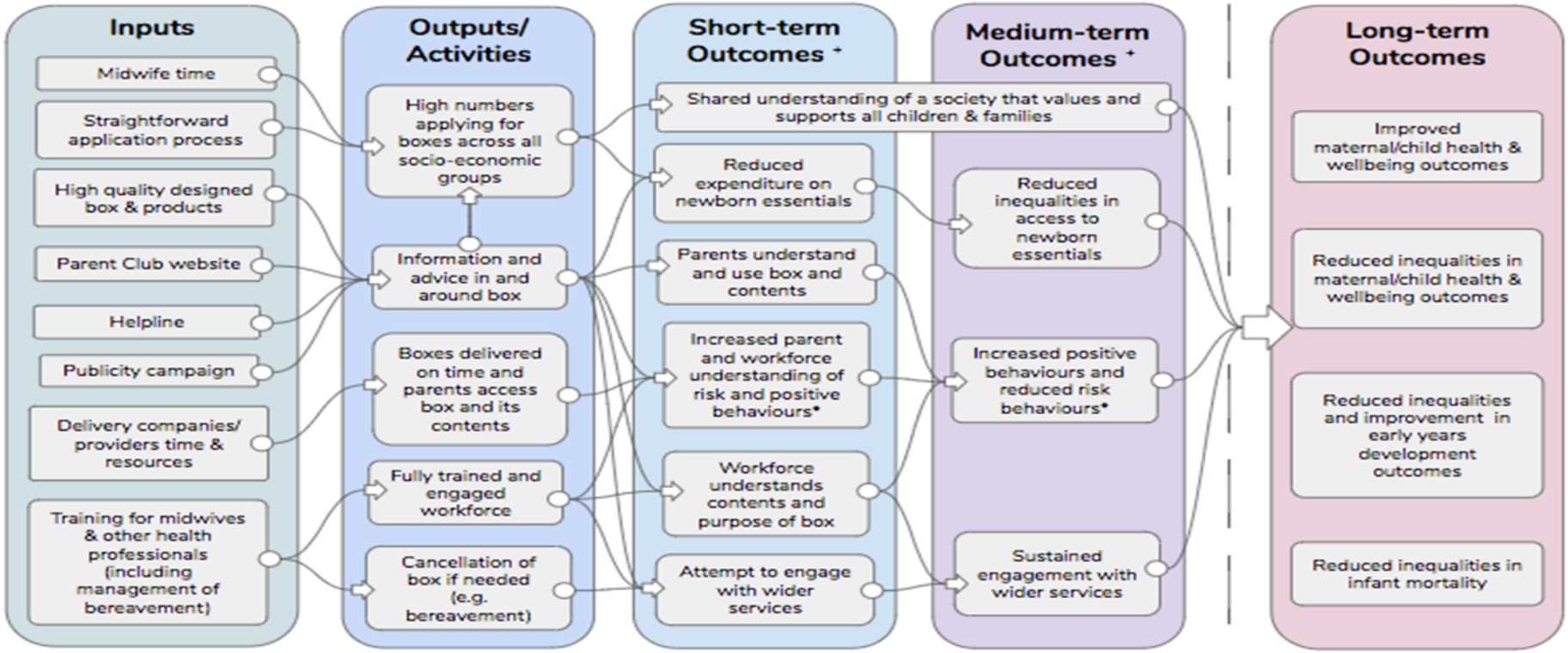
Stakeholders felt that training and education of the workforce to ensure that they understand the contents, purpose and use of the Box was a crucial short term outcome. Stakeholders also agreed that the initial outcome of “engagement with maternity services” should be changed to “attempt to engage with wider services,” to reflect the fact that the Baby Box may initiate engagement with services other than maternity services.

Stakeholders argued that seeking support was an important step in its own right, and should be measured whether or not support was obtained. Stakeholders also felt that the original long term outcome of “cultural shift in attitudes to equality – we all started life sleeping in a Box”, should be measured across the short and medium term, and

reworded as “shared understanding of a society that values and supports all children and families”.

Figure 1 shows the revised logic model, including the short term outcomes that were finally included.

Figure 1. Scotland's Baby Box – Logic model



Assumptions, synergies and risks:
Assumptions: Box adds to (rather than duplicating) parent's own resources (material and knowledge/information); Baby Boxes operating in wider package of maternity/family benefits – maternity grants, BSG, increased ELC provision.
Synergies: Operation of a wider package of maternity/family benefits – maternity grants, BSG, increased ELC provision.
Risks: Change in retailer behaviour (possible unintended consequences of universal provision of new born essential); media responses to Box (could undercut perceived usefulness/value and/or create stigma around accepting if not on board).
* Risk and Positive behaviours: breastfeeding, sleeping practice, health seeking behaviour and play/attachment.
+ The focus of the evaluation will be on the short and medium term outcomes, but the longterm outcomes should be passively monitored (as agreed upon by stakeholders at the first workshop on 24/1/2018).

Medium term outcomes

A key point raised during discussion was that the medium term outcomes should include reduction of specific risk behaviours as well as increasing positive behaviours as depicted by the original logic model. The medium term outcomes that stakeholders agreed to include in the revised logic model are shown in Figure 1.

Long term outcomes

The original logic model listed a number of longer term outcomes including lower infant mortality rates and reduced inequalities in infant mortality, improved parent child attachment and reduced inequalities in early learning outcomes. Participants noted that there were many other initiatives, such as the Universal Health Visiting Pathway and Best Start Grants, which might affect such outcomes.

It was also acknowledged that some relevant outcomes (such as serious adverse events) might be rare, or change only gradually following the introduction of the Baby Box. Following discussion, it was agreed that it would be difficult to attribute change in rare or longer term outcomes to the Baby Box and that the evaluation should focus on short to medium term outcomes, which could be more directly attributable to the Box.

It was also agreed that the focus of the evaluation should be on the Baby Box as a mechanism for engagement and education rather than as providing a safe sleeping space for infants. However, it was recognised that longer term child health, wellbeing and development outcomes should be monitored to identify possible impacts on health inequalities or other unintended consequences.

The final long term outcomes that stakeholders agreed should be monitored are outlined in Figure 1.

Workshop 2 (7 February 2018)

The main purpose of the second workshop was to identify potential primary and secondary data sources that could be used to measure short, medium and longterm outcomes.

Stakeholders considered the information collected during routine postnatal assessments in relation to the outcomes within the revised logic model. It was agreed that Health Visitors' assessment forms and other routinely collected data could be used to measure a range of short and medium term outcomes, but that additional primary data would be required to capture outcomes such as practitioners' and parents' understandings of the Baby Box and sustained engagement with services. In terms of longer term outcomes, such as reduced inequalities in infant mortality, it was recognised that National Records of Scotland (NRS) data could be used, supplemented by hospital admissions data to monitor other long term outcomes.

Workshop 3 (16 March 2018)

Workshop 3 involved identification and discussion of potential evaluation options, and assessment of their strengths and weaknesses.

It was noted that nationwide implementation of the Baby Box scheme ruled out a randomised trial of the scheme as a whole, though variations in provision could in principle still be trialled. Options considered included qualitative research with parents and practitioners, possibly supplemented by a quantitative survey, analysis of levels and trends in outcomes using routinely collected information to identify changes associated with the introduction of the Baby Box, and a prospective study of births in Scotland and a control area elsewhere in the UK, to identify differences in sleeping practices and other risk behaviours associated with receipt of the Box.

A number of constraints on evaluation design were noted. Only 25% of parents registered for a Baby Box provided consent to be contacted for research. This is not a problem for identifying a qualitative sample, but does create problems in using the database for survey sampling. There were also uncertainties about the suitability of the data recorded on the Child Health Systems Programme for use in an outcomes evaluation, for example where there are currently unreported elements being considered.

Evaluation options are set out and appraised in Chapter 3.

Chapter 3: Evaluation options

This chapter describes options for process and outcome evaluation identified during the workshops. Below each evaluation option, a table illustrates which outcomes would be addressed (tables 1-4). Pros and cons for each option are summarised in Table 5.

Option 0: No additional evaluation

Management information systems collect data on Baby Box registrations, including some information on applicants' demographic and geographical characteristics. No information is routinely collected through the registration process on outcomes or the mechanisms by which they are achieved. Without additional data gathering and analysis, evaluation would be limited to aspects of the registration and delivery process, such as analysis of regional or demographic variation in take up rates.

Option 1: Process evaluation

Stakeholders agreed that any approach adopted to evaluate the Baby Box should include a process evaluation in order to understand how the policy is implemented and how it may lead to outcomes.

Process evaluation recognises that intervention (in this case Scotland's Baby Box) outcomes depend on the interaction between the intervention and its context, and aims to understand the functioning of an intervention by examining implementation, mechanisms of impact, and contextual factors such as the characteristics of the target population and other services available to recipients of the intervention (Moore et al., 2015).

Understanding the role of context is particularly important for the evaluation of the Baby Box because other programmes and services, which also aim to improve health and wellbeing outcomes for children and families in Scotland, such as the Universal Health Visiting Pathway and Best Start Grant, are being introduced concurrently. Including a process evaluation alongside an outcomes evaluation can both enhance understanding of how outcomes are achieved, and contribute to improving design and implementation of the Baby Box.

A process evaluation can employ a range of data and methods, including qualitative interviews, focus groups, case note review, and quantitative surveys. Qualitative approaches use semi or unstructured interviews or focus groups, and samples that are designed to include participants with a range of relevant characteristics, rather than to be strictly representative of the whole population. Quantitative surveys use structured methods of data gathering, and representative samples.

1.1 Qualitative interviews and focus groups with parents and practitioners

Interviews and focus groups with parents and practitioners could be used to explore a range of short and medium term outcomes. For example, interviews and focus groups with parents could explore whether parents' understanding of what is important for their baby had been influenced by receiving the Baby Box, and whether they had consciously changed their plans or spending decisions in relation to new-born essentials.

They could also explore whether parents' awareness and understanding of risk and positive behaviours, such as breastfeeding and sleeping practices, had been influenced by receiving a Baby Box, whether it had encouraged them to engage with other services, and what kind of message(s) they thought the Baby Box projected as a feature of Government policy towards families with new babies.

Qualitative research of this kind will provide a deeper understanding of the context and processes behind particular impacts of the Baby Box but it will not provide generalisable information or give an indication of the prevalence of particular impacts.

Parents could be purposively sampled from those providing consent to be contacted for research. Although only 25% of parents registered agree to be contacted, the numbers should be sufficient to stratify the sample to ensure inclusion of a range of NHS Boards, different levels of the urban/rural classification and levels of deprivation. Further stratifying by the age of the baby would allow changes over time in sleeping practices, perceptions of the Box, and understanding of risk to be explored.

Interviews or focus groups with practitioners could be used to explore their understandings of the purpose of the Baby Box, how they perceived its influence on their interaction with parents, whether they felt able and confident to discuss the contents and purpose of the Baby Box with parents, and whether they felt it provided an opportunity to engage parents with other services. Sampling of practitioners should be stratified to match the sampling of parents so as to allow triangulation of the results of the parent and practitioner interviews.

Table 1 outlines the outcomes identified in the logic model that could be addressed by this option.

Table 1. Outcomes, example questions and methods for option 1.1

Short term outcomes	Examples of questions	Possible data sources
Reduced expenditure on new-born essentials	How much is the Baby Box saving families, especially low-income families, on new-born essentials?	Parent interviews or focus groups.
Parents understand and use Box and contents	Has the Baby Box changed parents' perceptions about what is essential for a baby,	Parent interviews or focus groups.

	including safety, positive interactions?	
Increased parents' and workforce understanding of risk and positive behaviours	Has the Baby Box improved parents' understanding of risk and positive behaviours, such as breastfeeding and sleeping practices? Has the Baby Box improved practitioners' understanding of risk and positive behaviours?	Parent interviews or focus groups. Midwife/Health visitor focus groups or interviews.
Workforce understands contents and purpose of box	Are practitioners confident, skilled and competent to discuss contents and purpose of Box with parents?	Midwife/Health visitor focus groups or interviews.
Attempt to engage with wider services	Have referrals to universal and wider services changed since the Baby Box? Is the Baby Box offering new opportunities for practitioners to engage with parents stemming from the information pack provided in the Box? Is the Baby Box offering new opportunities to identify families who are unlikely to engage with services?	Practitioner interviews or focus groups with practitioners. Case note review.
Medium term outcomes	**Examples of questions**	**Possible data sources**
Shared understanding of a society that values and supports all children	Does the Baby Box change people's perceptions about universal benefits? Has uptake of the Baby Box changed over time; is this a reflection of a cultural change?	Parent interviews or focus groups. Parent interviews or focus groups. Case note review.
Increased positive behaviors and reduced risk behaviors	Are parents increasingly engaging in positive behaviours and reducing risk behaviours?	Parent interviews or focus groups.
Sustained engagement with wider services	Is there greater access to and uptake of services, where information is provided, to improve health and wellbeing for parents?	Health Visitor interviews or focus groups. Case note review. Practitioners in other services interviews or focus groups.

1.2 Quantitative surveys of parents and practitioners

(a) A survey of parents sampled from the registration database (of those that have consented to further contact for research) could be conducted to examine short and medium term outcomes, such as expenditure on new-born essentials, the extent to which particular items in the Box are used and the prevalence of risk or protective behaviours. This will provide quantitative information that is generalisable to the wider population. As noted, the proportion of parents who currently consent to research is about 25%. As those who consent are unlikely to be representative of all parents registered for the Baby Box, it will be necessary to stratify the sample to ensure that some groups (for example more affluent or better educated parents) are not over-represented. Disproportionate sampling and any additional imbalance in response could be corrected by applying appropriate weights during analysis, with some loss of statistical efficiency.

(b) Scottish Government officials mentioned that formal training on the Baby Box for health professionals could be considered in the future. A baseline survey could be conducted prior to the introduction of a formal training programme. Once the training has been implemented, existing reporting mechanisms could be used to monitor participation and a further survey carried out with practitioners who have received the training to ascertain whether their knowledge, skills and competencies to discuss the purpose, contents and use of the Box with parents had improved.

Table 2 outlines the outcomes identified in the logic model that could be addressed by this option.

Table 2. Outcomes, example questions and methods for option 1.2

Short term outcomes	Examples of questions	Possible data sources
Reduced expenditure on new-born essentials	How much is the Baby Box saving families, especially low-income families on new-born essentials?	Surveys of parents.
Increased parents' and workforce understanding of risk and positive behaviours	Has the Baby Box improved parents' understanding of risk and positive behaviours, such as breastfeeding and sleeping practices? Has the Baby Box improved practitioners understanding of risk and positive behaviours?	Surveys of parents. Midwives'/Health visitors' self-reported knowledge, skills and competencies – surveys.
Workforce understands contents and purpose of Box	Are practitioners confident, skilled and competent to discuss contents and purpose of Box with parents?	Midwives'/Health Visitors' self-reported knowledge, skills and competencies – surveys.
Attempt to engage with wider services	Have referrals to universal and wider services changed since the Baby Box?	Midwives'/Health Visitors' self-reported knowledge, skills and competencies-

	Is the Baby Box offering new opportunities for practitioners to engage with parents? Is the Baby Box offering new opportunities to identify families who are unlikely to engage with services?	surveys.
Medium term outcomes	**Examples of questions**	**Possible data sources**
Shared understanding of a society that values and supports all children	Does the Baby Box change peoples' perceptions about universal benefits? Has uptake of the Baby Box changed over time; is this a reflection of a cultural change?	Surveys with parents.
Reduced inequalities in access to new-born essentials	Is there improved and more equitable access to new-born essentials?	Survey of parents (comparing across SES groups).
Increased positive behaviors and reduced risk behaviors	Are parents increasingly engaging in positive behaviours and reducing risk behaviours?	Surveys of parents.
Sustained engagement with wider services	Is there greater access to and uptake of services to improve health and wellbeing for parents?	Surveys of parents.

Option 2: Outcome evaluation using routinely collected data

The Baby Box is expected to achieve a wide range of outcomes, over a range of timescales, making evaluation of its impact using routinely collected data challenging. The Baby Box has already been implemented on a universal basis across the whole of Scotland, and so an experimental trial of the scheme as a whole is no longer a feasible or ethical option. We focus in this section on approaches that can be applied retrospectively using routinely collected data. Under option 3 below we consider a research design that involves prospective data gathering but in the context of an observational rather than experimental study design.

As noted above, stakeholders agreed that the focus for any outcome evaluation should be on short or medium term outcomes, given the difficulty of attributing longer term outcomes to the Baby Box in the context of other interventions for babies and parents such as the Best Start Grant and the Universal Health Visiting Pathway. It was also agreed that some rare or longer term outcomes related to child and maternal health, including sudden unexpected death in infancy, should be monitored only (rather than being part of the evaluation) to check that there were no adverse changes in levels, trends or socio-economic patterning of these outcomes that might require further scrutiny. Option 2 therefore considers how routinely collected data

could be used to measure change in short/medium term outcomes, such as sleeping practices, development outcomes, referrals to other services, etc.

One possible approach to measuring the impact of the Baby Box would be to use interrupted time series methods to analyse changes in the level and trend of routinely monitored child health outcomes before and after the introduction of the Baby Box. Information on the health of all babies and young children in Scotland is collected via the Child Health Systems Programme (CHSP) Pre-School system of assessments. The present system has largely been in place since 2013, though the forms currently in use were introduced in 2016, with further changes in April 2017 including the introduction of additional assessments at 13-15 months and 4-5 years as part of the Universal Health Visitor Pathway. Health Visitors complete standardised forms at the first post-partum visit around 10 days following a birth, 6-8 weeks post birth, and when the child is 13-15 months, 27-30 months and 4-5 years.

The assessments collect a range of information relevant to the evaluation of the Baby Box. At the first post-partum visit, information is collected on smokers in the household and infant feeding (at birth, hospital discharge and current method). The 6-8 week review collects information on feeding (breast, bottle or both); parental concerns (feeding, appearance, behaviour, hearing, eyes, sleeping, movement, illness, crying, weight gain, and other); development (gross motor, hearing and communication, vision and social awareness); physical measures; diagnoses and concerns related to the child's health, development and wellbeing; sleeping position (prone, supine and side); and referrals to other services. Subsequent assessments collect information on feeding (13-15 month only); development; physical measures; diagnoses/concerns related to the child's health, development and wellbeing; and referrals to other services.

CHSP Pre-School system of assessments information is a potentially valuable resource for evaluating the effect of the Baby Box on infant feeding, sleeping position, development, and parents' engagement with services but there are significant limitations that need to be taken into account. Implementation is the responsibility of NHS Boards, and practice varies within and between Boards. Although national guidance for the 27-30 month assessment was issued in April 2017 recommending that all Boards use the same development assessment tool (the Ages and Stages Questionnaire, version 3), practice before then differed widely (ISD 2018a, ISD n.d.). ISD note that 'in some NHS Boards a phased implementation occurred after the system was adopted. Therefore, caution should be taken when interpreting data around the implementation period. In addition, many data items are not mandatory and recording practices vary between NHS Boards and individual health professionals. Therefore, not all data items on the review forms can be used for analysis (ISD 2018b).

While some of the information, particularly the items used in national reporting such as breastfeeding, is believed to be of high quality (ISD 2017), other items (including sleeping position) that are not routinely analysed have not been validated, and

validation work would be needed to ascertain whether they could be used for evaluation purposes.

The child development data is potentially useful but there may be a need to stratify analyses according to the tools used, or to focus on a subset of Boards using particular development assessment tools over a period of years. Implementation of the 13-15 month assessment overlapped with implementation of the Baby Box, so only Boards that introduced the additional assessment rapidly could provide data both pre and post implementation of the Baby Box, and the length of the pre-implementation series would be limited.

If a consistent monthly series of data could be obtained from the first post-partum visit and 6-8 week assessments for relevant outcomes such as feeding, sleeping position and development, interrupted time series methods could be used to identify changes in the levels and/or trends in those outcomes associated with the introduction of the Baby Box. This method would control for pre-intervention trends in the outcomes of interest, but not for the effects of other events that occur around the same time as the Baby Box was implemented, such as other changes in ante- or post-natal care.

Interrupted time series approaches are well-suited to evaluating policies that are implemented at a specific point in time, such as the Baby Box. The analyses could use aggregate data (e.g. monthly counts) rather than individual data, and would not require linkage of Baby Box registrations with the CHSP Pre-School system of assessments data. However, it would be necessary to use individual level data for analyses stratified by parental characteristics such as age and deprivation scores. Without linkage to application data, the analyses would provide ‘intention to treat’ (ITT) estimates, identifying the effect of the Baby Box on all those eligible, rather than the effect on recipients.

As the take-up of the Baby Box is high (83% in June 2018) but not complete, the ITT and ‘on treatment’ analyses may differ. Similar methods could eventually be applied to the 27-30 month assessment data but attributing changes in outcomes at that stage to the Baby Box would be problematic for the reasons noted above, and a post-intervention series of data would not be available until late 2019.

Table 3 outlines the outcomes identified in the logic model that could be addressed by this option.

Table 3. Outcomes, example questions and methods for option 2

Short term outcomes	**Examples of questions**	**Possible data sources**
Attempt to engage with wider services	Have referrals to universal and wider services changed since the Baby Box? Is the Baby Box offering new opportunities for practitioners to engage with parents? Is the Baby Box offering new opportunities to identify families who are unlikely to engage with services?	ISD data – Health Visitors' records of first visit, 6-8 weeks visit. Data from other services if available.
Medium term outcomes	**Examples of questions**	**Possible data sources**
Increased positive behaviors and reduced risk behaviors	Are parents increasingly engaging in positive behaviours and reducing risk behaviours?	ISD data – Health Visitors' records.
Sustained engagement with wider services	Is there greater access to and uptake of services to improve health and wellbeing for parents?	ISD data – Health Visitors' records.
Long term outcomes	**Examples of questions**	**Possible data sources**
Improved maternal/child health and wellbeing outcomes	Does exposure to second hand smoke change after the introduction of Baby Box? To be monitored over time to track changes.	ISD data – Health Visitors' records. ISD data – hospital admissions.
Reduced inequalities in maternal/child health and wellbeing outcomes	Is there a reduction in inequalities in maternal/child health after introduction of the Baby Box? To be monitored over time to track changes in health inequalities.	ISD data – Health Visitors' records.
Reduced inequalities and improvement in early years development outcomes	Is there a change in inequalities and improvement in early years development outcomes after the introduction of the Baby Box? To be monitored over time to track changes in health inequalities.	ISD data – Health Visitors' records.
Reduced inequalities in infant mortality	Is there a change in inequalities in infant mortality after the introduction of the Baby Box? To be monitored over time to track changes in in infant mortality.	National Records of Scotland data.

Option 3: Birth Cohort Study

An alternative to the use of routinely collected data that was suggested at the third workshop was to collect primary data from a prospective cohort of births in Scotland and a comparison area (or areas) in England or Wales. This would allow a much finer-grained assessment of sleeping practices and other key outcome domains, for example using diaries kept by participants. It would provide information about how practices differed between recipients of the Baby Box and non-recipients. Note however that such differences could not be interpreted as effects of the Baby Box because there would be no information on how practices had changed following its introduction, and the effects of the Baby Box would be confounded by other differences between Scotland and the comparison area(s) in services provided.

If the sample were stratified, for example by deprivation scores based on participants' postcodes, a cohort study could also provide insights into socio-economic variation in sleeping practices and other key behaviours among recipients and non-recipients. As with the overall impact of the Baby Box, it would not be possible to estimate impacts of the Box on inequalities because information about changes in the socio-economic patterning of outcomes following the introduction of the Box would be lacking.

A cohort study of this kind would be a very substantial undertaking. The advantages in terms of flexibility to determine what data should be gathered need to be weighed against the costs and risks involved in attempting to recruit participants during pregnancy or very shortly after birth, and collecting substantial amounts of data in the early stages of parenthood. Extensive development, consideration of ethics, feasibility and piloting work would be needed to: determine the best way to engage maternity services to manage the identification and recruitment of parents; approach parents (or prospective parents) to take part in the study; estimate required sample sizes; and identify (or develop) valid and reliable instruments for gathering data. It would be substantially more costly and time-consuming than a partly or wholly retrospective study relying on routinely collected data but may be worth considering if other monitoring or research raised concerns about possible adverse effects associated with the Baby Box.

Table 4 outlines the outcomes identified in the logic model that could be addressed by this option.

Table 4. Outcomes, example questions and methods for option 3

Medium term outcomes	Examples of questions	Possible data sources
Reduced inequalities in access to new-born essentials	What are the differences in access to new born essentials in the Baby Box and non-Baby Box area?	Cohort survey of parents (comparing across SES groups).
Increased positive behaviors and reduced risk behaviors	Are parents increasingly engaging in positive behaviours and reducing risk behaviours?	Cohort survey of parents (including sleep and feeding diaries).

Table 5 summarises the evaluation options.

Table 5. Summary of evaluation options

Option	Description	Comments
Option 0	No additional evaluation	Management information systems will provide some data on inputs and outputs, such as volume of applications, take-up rates, etc., but no information on outcomes or the mechanisms by which they are achieved.
Option 1.1 – Process - qualitative	Qualitative process evaluation based on interviews and/or focus groups with purposive sample of practitioners and recipients; recipient sample to be drawn from those providing consent to contact for research in their application	Will provide information on process of delivery, mechanisms of impact and contextual factors (e.g. receipt of other interventions), and insights into the way the Baby Box is perceived and used by practitioners and recipients. For example, short term outcomes 'has the Baby Box improved parents' understanding of risk and positive behaviours'. Will not provide quantified estimates of outputs or short/medium term outcomes, or any information about longer term outcomes.
Option 1.2 – Process - quantitative	(a) Survey (by phone or face-to-face) of recipients sampled from applicant database (b) Surveys of practitioners	Would provide quantitative estimates of short/medium term outcomes, such as expenditures on new-born essentials, levels of understanding of risk/protective behaviours, etc., but would not allow identification of impact due to lack of control/counterfactual. Parents who consent to research (25% of total) are unlikely to be representative of all applicants; an opt-in procedure prior to interview may accentuate this bias. It could be corrected with post stratification weights, with some loss of statistical efficiency. Surveys of practitioners before and after training could be used to assess and improve the effectiveness of training. Can be hard to achieve good response rates.
Option 2 – Outcomes evaluation using retrospective routine data - quantitative	Quantitative evaluation of routinely collected data from CHPS Pre-School data	Would provide quantitative estimates of the impact of the Baby Box on some short/medium term outcomes, such as levels and trends of sleeping practices, development outcomes, play/attachment, referrals to other services, etc., using interrupted time series to control for pre-intervention trends. Consistent series of data only likely to be available for outcomes recorded at first visit or 6-8 weeks.

		Highly dependent on quality/coverage of routinely collected data, which is unknown for many outcomes. Would identify the effect of the Box on all those eligible, rather than on recipients.
Option 3 – Outcomes evaluation using prospective cohort	Prospective cohort of births in Scotland and England/Wales	Could provide detailed information, specific to the requirements of the evaluation, on a range of short/medium term outcomes, such as sleeping practices and other risk/protective behaviours, and quantitative estimates of differences in these behaviours between Baby Box recipients and non-recipients. Would not provide direct estimates of the impact of the Baby Box. Methodologically challenging, and would require extensive development, feasibility and piloting work. It would also need to go through a medical ethical review process.

Chapter 4: Discussion

The range of questions identified as priorities in the workshops indicates a mixed methods approach to evaluation, combining primarily qualitative investigation of the process of delivery from the perspective of both parents and practitioners, parents' use of the Baby Box, and their perceptions of its meaning and value, with primarily quantitative investigation of the impact of the Baby Box on parenting behaviours and child health outcomes.

We have not explored options for economic evaluation of the Baby Box, as this was not identified as a priority in the workshops. A cost-effectiveness analysis would require the identification of a primary outcome (or a small set of outcomes) which could be used to compare the Baby Box against an alternative course of action. This may not be appropriate given the wide range of aims identified for the Baby Box. A cost consequence analysis, which simply presents the whole array of outputs alongside the costs, may be preferable if an economic evaluation is required.

The consent for research included within the Baby Box registration card provides a potentially valuable register for obtaining samples of recipients but, as it has not yet been used for evaluation, further investigation of the social patterning of consent should be carried out before it can be relied on to obtain representative samples. Likewise, the CHPS Pre-School system opens up the possibility of using routinely collected data to measure outcomes at much lower cost than approaches based on prospective, primary data collection but extensive validation work is needed to establish which outcomes, if any, could usefully be captured from this source.

These considerations, together with the fact that there have been no previous evaluations of Baby Box interventions delivered on a whole population basis, suggest a phased approach to commissioning evaluation may be advisable, with a substantial initial phase of feasibility testing and development work prior to final decisions about the design of a substantive process and outcome evaluation.

The feasibility/development phase might include: analysis of the pattern of consent/non-consent to take part in research among Baby Box registrations according to socio-demographic characteristics such as age, birth parity or SIMD; analysis of the quality of data on outcomes, such as sleeping position, child development, referrals to other services, etc., recorded in the CHSP Pre-School system of assessments; comparison of CHSP Pre-School with systems used elsewhere in the UK for recording early years outcomes, in order to identify possibilities for using geographical controls in the outcome analyses, as well as controls for trends in outcomes over time; and scoping reviews of previous research on the impacts of benefits in kind provided during pregnancy in developed countries, and on methodologies (such as diaries) for collecting detailed information on safe sleeping and other parenting practices that might be influenced by the Baby Box.

The approaches we have discussed for identifying the effects of the Baby Box are observational rather than experimental, in that they would rely on data collected

alongside the implementation of the policy rather than randomly assigning some people to receive the Box, and comparing their outcomes with those of people who do not receive the Box.

Observational methods are widely used to evaluate public health and other policies where random assignment is impractical or considered unethical but they are subject to a number of limitations in relation to the kinds of effects that can be measured and the extent to which observed differences in outcomes could be attributed to the Baby Box. These limitations could in principle be overcome by a randomised controlled trial.

While a randomised control trial of the Baby Box scheme as a whole would be impractical in the context of full implementation, it should be noted that experimental trials of variants of the Baby Box (e.g. with additional items or information) or alternative modes of delivery (e.g. with additional professional/practitioner input for some or all recipients) would not necessarily be unethical or impractical. Trials may be worth considering if the evaluation identifies areas for improvement.

APPENDICES

Appendix 1. Stakeholders who attended at least one workshop

Name	Organisation	Role
Evaluability Assessment Collaborative		
Peter Craig	MRC/CSO Social & Public Health Sciences Unit, University of Glasgow and What Works Scotland	Academic Facilitator
Larry Doi	Scottish Collaboration for Public Health Research and Policy, University of Edinburgh	Academic Facilitator
Laura Tirman	Scottish Collaboration for Public Health Research and Policy, University of Edinburgh	Academic Facilitator
Stakeholders		
Alan Davidson	Scottish Government	Official
Andrea Kinver	Scottish Government	Official
Anthea Taylor	Scottish Government	Official
Beverly Walsh	APS Group	Head of Business Services, APS Group Scotland
Claire Jamieson	Scottish Government	Official
Clare Simpson	Parenting Across Scotland	Manager
Dave Gorman	Scottish Government	Official
Dorothy Ann Timoney	NHS Lothian	Family Nurse Partnership Nurse
Evelyn Frame	NHS Greater Glasgow and Clyde/ Scottish Government	Chief Midwife/Midwife Advisor
Franca Macleod	Scottish Government	Official
Helen Ball	Durham University	Academic
Helen Crosthwaite	APS Group Scotland	Client Services Director
Jennifer Bonnar	NHS Lothian	Family Nurse Partnership Nurse
Joanna Shedden	Scottish Government	Official
Kim Wallace	Scottish Government	Official
Lynne Miller	Family Therapy Training Network	Family Therapist
Lynsay Allan	Cot Death Trust	Director
Mary Ann Gillian	NHS Fife	Midwife
Nicola Welsh	Sands Lothians	Director
Pam Murray	NHS Lothian	Family Nurse Partnership Nurse
Pete Blair	University of Bristol	Academic
Ruari Sutherland	Scottish Government	Official
Sam Pringle	Father's Network	Director
Shelagh Considine	Scottish Government	Official
Shelagh Young	Homestart UK	Director of Scotland
Susan Gallacher	Scottish Government	Official
Una MacFadyen	Forth Valley Royal Hospital	Consultant Paediatrician

Appendix 2. References

Baddock SA, Tipene-Leach D, Williams SM et al (2017) Wahakura Versus Bassinet for Safe Infant Sleep: A Randomized Trial. Pediatrics, 139(2). pii: e2016016.

Embassy of Finland, Washington (2017) Finnish Baby Box Inspires around the World. http://www.finland.org/public/default.aspx?contentid=357875&culture=en-US (last accessed 7 June 2018).

Gissler M and Hakulinen T (2017) Finland's low infant mortality has multiple contributing factors. https://blogi.thl.fi/finlands-low-infant-mortality-has-multiple-contributing-factors/ (last accessed 7 June, 2018).

ISD (2017) Infant Feeding Statistics Scotland Technical Report Financial Year 2016/17.

ISD (2018a) Child Health 27-30 Month Review Statistics Technical Report Scotland 2016/17.

ISD (2018b) Child Health Systems Programme Pre-School System of Assessments (CHSP Pre-School) http://www.isdscotland.org/Health-Topics/Child-Health/Child-Health-Programme/Child-Health-Systems-Programme-Pre-School.asp#users (last accessed, 25 April 2018).

ISD (n.d.) Use of developmental assessment questionnaires within 27 month child health reviews: insights from the first year of review delivery. http://www.isdscotland.org/Health-Topics/Child-Health/Child-Health-Programme/docs/Use-of-developmental-assessment-questionnaires-within-27-month-child-health-reviews-final.pdf (last accessed, 25 April 2018).

Moore G, Audrey S, Barker M et al. (2015) Process evaluation of complex interventions. Medical Research Council guidance. *BMJ* 2015;350:h1258 doi:10.1136/bmj.h1258.

OECD (2018), Infant mortality rates (indicator). doi: 10.1787/83dea506-en (Last accessed, 7 June 2018).

Tipene-Leach D, Baddock S, Williams S et al. (2014) Methodology and recruitment for a randomised controlled trial to evaluate the safety of Wahakura for infant bed sharing. BMC Pediatrics. 2014 Sep 28;14:240. doi: 10.1186/1471-2431-14-240.

Wood R, Stirling A, Nolan C, et al. (2012) Trends in the coverage of 'universal' child health reviews: observational study using routinely available data. BMJ Open;2:e000759. doi:10.1136/bmjopen-2011-000759.

social research

The views expressed in this report are those of the researcher and do not necessarily represent those of the Scottish Government or Scottish Ministers.

This document is also available from our website at www.gov.scot.
ISBN: 978-1-78781-242-0

The Scottish Government
St Andrew's House
Edinburgh
EH1 3DG

Produced for
the Scottish Government
by APS Group Scotland
PPDAS470686 (09/18)
Published by
the Scottish Government,
September 2018

GSR
GOVERNMENT SOCIAL RESEARCH
Social Science in Government

Social Research series
ISSN 2045-6964
ISBN 978-1-78781-242-0

Web and Print Publication
www.gov.scot/socialresearch

PPDAS470686 (09/18)